Chemical Variation in a Dominant Tree Species: Population Divergence, Selection and Genetic Stability across Environments

Julianne M. O’Reilly-Wapstra1,3*, Alison M. Miller1,3, Matthew G. Hamilton1,3, Dean Williams2, Naomi Glancy-Dean1, Brad M. Potts1,3

1 School of Plant Science, University of Tasmania, Hobart, Tasmania, Australia, 2 Forestry Tasmania, Hobart, Tasmania, Australia, 3 National Centre for Future Forest Industries, University of Tasmania, Hobart, Tasmania, Australia

Abstract

Understanding among and within population genetic variation of ecologically important plant traits provides insight into the potential evolutionary processes affecting those traits. The strength and consistency of selection driving variability in traits would be affected by plasticity in differences among genotypes across environments (G×E). We investigated population divergence, selection and environmental plasticity of foliar plant secondary metabolites (PSMs) in a dominant tree species, Eucalyptus globulus. Using two common garden trials we examined variation in PSMs at multiple genetic scales; among 12 populations covering the full geographic range of the species and among up to 60 families within populations. Significant genetic variation in the expression of many PSMs resides both among and within populations of E. globulus with moderate (e.g., sideroxylonal A $h^2_{op} = 0.25$) to high (e.g., macrocarpal G $h^2_{op} = 0.48$) narrow sense heritabilities and high coefficients of additive genetic variation estimated for some compounds. A comparison of $Q_{st}$ and $F_{st}$ estimates suggest that variability in some of these traits may be due to selection. Importantly, there was no genetic by environment interaction in the expression of any of the quantitative chemical traits despite often significant site effects. These results provide evidence that natural selection has contributed to population divergence in PSMs in E. globulus, and identifies the formylated phloroglucinol compounds (particularly sideroxylonal) and a dominant oil, 1,8-cineole, as candidates for traits whose genetic architecture has been shaped by divergent selection. Additionally, as the genetic differences in these PSMs that influence community phenotypes is stable across environments, the role of plant genotype in structuring communities is strengthened and these genotypic differences may be relatively stable under global environmental changes.

Introduction

Phenotypic variation in traits across the geographic range of a species can be the result of a variety of environmental effects and/or genetic processes such as local adaptation and drift [1,2]. A key interest for many ecologists and evolutionary biologists is to unravel these genetic processes to better understand evolutionary forces affecting ecologically relevant traits [3]. A useful approach to achieve this is to investigate how phenotypic variation is partitioned at different genetic levels, such as among and within populations [4,5,6]. Among population variation in traits may indicate population divergence in trait evolution while estimation of within population genetic parameters can be used to understand how these traits may respond to selection. For example, narrow sense heritability is a useful within population genetic parameter to estimate the proportion of phenotypic variance that is due to additive effects (and thus is the variance selection can act upon). This matched with the coefficient of additive genetic variance (the additive genetic variance standardised by the trait mean) to assess the ‘evolvability’ of a trait, provides a clear methodology for understanding the evolutionary potential of traits [7,8].

Plant/herbivore systems represent a model approach to investigate patterns of evolutionary population divergence across the geographic range of a species and the evolutionary processes affecting relevant traits within populations. Major drivers of plant/herbivore interactions are plant secondary metabolites (PSMs; compounds produced as a result of plant secondary metabolism) [9,10,11,12]. Many studies have investigated the evolutionary processes affecting the expression of PSMs in plants [13,14,15], however, relatively few have taken a hierarchical genetic approach to better understand the genetic basis of natural variation in these compounds [5]. In addition, the strength and consistency of evolutionary pressures in driving variability in PSMs would be affected by variation in the plasticity of genotypes in the way they respond to environmental variation (G×E) [16], and these traits may also be affected by correlated responses to selection [17]. Consequently, studies that take a hierarchical genetic approach, combined with genotype by environment manipulations and an understanding of how multiple PSM traits are genetically...
correlated, are beneficial to understand the multiple factors that lead to variation in the evolution of PSMs in plant/herbivore systems.

*Eucalyptus globulus* is a dominant tree species of ecological significance in south-eastern Australia, and is rich in foliar PSMs including terpenes and phenolic based compounds such as formylated phloroglucinol compounds (FPCs). These compounds affect rates of herbivory by many marsupial herbivores and structure foliar arthropod communities [18,19,20,21]. The quantitative genetic expression of some of these traits in *E. globulus* is well understood, however, previous work has focused on among population genetic variation [18,19,22]. Little is known about the levels of within population level genetic variation (and hence ‘evolvability’) of these key plant traits. Even less is known about how variable genotypes are in the expression of PSMs at both the within and among population levels across different environments [23] and how these chemical traits relate to growth or fitness traits in the species [24].

We advance previous research on broad-scale geographic genetic variation in PSMs in *E. globulus* [22,25] by taking a within population, family (sibling) level approach to better understand genetic variation in key PSMs across multiple genetic scales (by comparing among and within population variation) and to better understand variation in these traits at the scale at which natural selection operates (within populations). Due to our addition of family pedigrees we can estimate quantitative genetic parameters including narrow sense heritability, coefficient of additive genetic variation and compare estimates of quantitative (Qnt) versus neutral (Fnt) genetic variation. All of these estimates provide important information on a traits ability to respond to selection [1,7,26]. We also examine the inter-trait correlations between these compounds across the different levels of the genetic hierarchy (population and family) to investigate if within population genetic correlations relate to among population patterns of divergence in these traits. We examine how these traits relate to plant growth and browsing damage in the field and importantly, we examine the genetic-based plasticity of PSMs at these different genetic scales, across different environments.

**Materials and Methods**

**Ethics statement**

No permits or approvals were required for this field study as only small samples of an abundant tree species were taken. Effects to individual and ecosystem health were negligible. The land accessed is not protected. No protected species were sampled.

**Genetic pedigree of *E. globulus***

*Eucalyptus globulus* has been classified into a genetic hierarchy of 13 geographically and genetically distinct races and 20 sub-races across its full geographic range, based on differences in 35 different morphological traits assessed in a range of common environment trials [27]. Families within sub-races are offspring from a single mother tree [27,28].

**Trial design, sampling and browse assessment**

In this study we utilised two common environment field trials established with progeny of known family pedigree and representative of a large proportion of *E. globulus* sub-races. Foliage was collected from these progeny trials located at Togari (40° 57’ S, 144° 55’ E, 90 m asl) and Salmon River (41° 02’ S, 144° 49’ E, 100 m asl) in north-west Tasmania. Both field trials were established in August 2005 by Forestry Tasmania. Sites differed markedly in their soil profiles. Salmon River was yellow brown mottled clay on Precambrian mudstone whereas Togari was red brown clay on Cambrian inter-layered mudstone, siltstone and sandstone. Each trial comprised plants grown from open-pollinated seed collected from 310 (Salmon River) and 244 (Togari) families collected from throughout the native, natural distribution of *E. globulus* (see Dutkowski & Potts [27] for details). Furthermore, seedlots from breeding programmes and commercial seed-orchards were included in the trials. Each trial was a randomised incomplete block design with each family represented as a single-tree plot. Salmon River (SR) comprised 20 replicates of 15 incomplete blocks per replicate while Togari (TO) comprised 16 replicates of 12 incomplete blocks.

Juvenile foliage was harvested from a subset of trees at each site in March 2006. Two healthy leaves were taken from each individual tree. Foliage was selected from the northeast side, half way up the height of the seedlings (seedlings were approximately one metre tall), three leaf pairs back from branch tips. This assured that fully expanded juvenile leaves (but not yet entering senescence) were sampled and that leaves sampled were the same age across all plants. Trees harvested from Salmon River represented 60 families from 12 sub-races while samples were taken from trees representing 53 families from the same 12 sub-races at Togari. These were collected across 20 and 15 replicates from Salmon River and Togari, respectively. In total 749 samples were collected from the two sites (Table 1). Sub-race locations are shown in Figure 1. Leaves were placed in cool storage and frozen on return to the laboratory. Leaves were later freeze-dried in preparation for predictive chemical analysis using near infrared reflectance spectroscopy, after which time they were ground in a cyclone mill for wet chemical analysis.

Mammal browsing damage was scored at Togari in March 2006. Trees at Salmon River River were not browsed. Browsing damage was scored as percentage of total foliage removed from the whole plant on a scale from 0 to 6, where 0 = 0%, 1 = 1–5%, 2 = 6–25%, 3 = 26–50%, 4 = 51–75%, 5 = 76–95%, and 6 = 96–100%. Due to generally low browsing (~27% of trees received damage), scores were converted to a binary presence/absence for analysis. Diameter at breast height (DBH; 1.3m) was assessed in February 2010 after 5 years growth to relate browsing and foliar chemical variation with a surrogate measure of tree fitness.

**Chemical analysis**

**Near infrared reflectance spectroscopy.** We predicted the concentration of six different chemical constituents in the leaves using near infrared reflectance spectroscopy (NIRS). These constituents were nitrogen, two formylated phloroglucinol compounds (sideroxylonal A and macrocarpal G), total oils, 1,8-cineole (the dominant oil in *E. globulus*) and total phenolics. Near infrared reflectance spectroscopy has been extensively in ecology and other fields to enable rapid analysis of samples on a scale that would be impractical using conventional laboratory techniques (see Foley et al. [29] for a review). For each sample (n = 749) two freeze-dried leaves were scanned with near-infrared light using a Bruker MPA FT-NIR spectrometer coupled to a fibre-optic probe. The tip and the basal region of the adaxial surface of each leaf would be impractical using conventional laboratory techniques (see Foley et al. [29] for a review). For each sample (n = 749) two freeze-dried leaves were scanned with near-infrared light using a Bruker MPA FT-NIR spectrometer coupled to a fibre-optic probe. The tip and the basal region of the adaxial surface of each leaf were scanned, avoiding the midrib and damaged areas. Spectra (expressed as log (1/R)) were collected from 7690–4350 cm⁻¹ (1300–2300nm) at a resolution of 4 cm⁻¹. The resulting four spectra taken from the two leaves were averaged.

**Near infrared reflectance spectroscopy model development and predictions.** Near infrared reflectance spectroscopy models were developed to predict the concentration of all chemical constituents. Prior to model development, principal
component analysis (PCA) was carried out on the spectra of all samples, to enable selection of a subset (n = 100; 50 samples per site) of samples for wet chemical analysis (see below) for model validation purposes. These selected samples were representative of the spectral variation present in the entire sample set. Half of these samples (n = 50) were added to existing NIRS models [30] and these models were re-developed to predict the concentration of the samples used in this study. The other half of the samples (n = 50) were used as an independent validation set of the predictive models. The oil models were developed using only the samples used for wet chemical analysis in this current study (that is, not built from existing models). All models were developed using internal cross validation to avoid overfitting. Models showed a close relationship between wet chemistry and predicted values, both within the calibration set and upon independent validation. Internal cross validation results for the NIRS models ranged from r² 0.66–0.90 (SEP: 0.22–5.21) for the six chemical traits; the external validation r² ranged from 0.69–0.86 (SEP: 1.29–5.05).

Wet chemistry. Wet chemical analysis was carried out on 50 samples per site. Total oils and 1,8-cineole were assayed by gas chromatography-mass spectrometry (GCMS). The method was modified from that reported in O’Reilly-Wapstra et al. [19] whereby freeze-dried ground samples were used for extraction and results were expressed as mg g DM⁻¹ equivalents of cineole using a 1,8-cineole standard. Total phenolics were assayed with a modified Prussian blue assay for total phenolics using gallic acid standards [31] after extraction following the method outlined in Hagerman [32]. Using a gallic acid standard, total phenolics concentration was expressed as mg g DM⁻¹ equivalents of gallic acid. Sideroxylonal A and macrocarpal G were assayed by high performance liquid chromatography (HPLC) following Wallis & Foley [33]. By using a pure standard of sideroxylonal A, results were expressed as mg g DM⁻¹. Results for macrocarpal G concentration were expressed as mg g DM⁻¹ equivalents of macrocarpal A (using a macrocarpal A standard). Freeze dried ground leaf samples were then ground to a fine powder using a ball grinder (Retsch MM200) and analysed for nitrogen using an elemental analyser (Thermo Finnigan EA 1112 Series).

### Table 1. Number of families and individuals sampled at each hierarchical genetic scale.

<table>
<thead>
<tr>
<th>Sub-race</th>
<th>Sub-race code</th>
<th>Families</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SR</td>
<td>TO</td>
</tr>
<tr>
<td>Eastern Otways</td>
<td>EOt</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Flinders Island</td>
<td>FI</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Southern Furneaux</td>
<td>SF</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>King Island</td>
<td>Ki</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>St Helens</td>
<td>STH</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>North-eastern Tasmania</td>
<td>NET</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Recherche Bay</td>
<td>RB</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>South-eastern Tasmania</td>
<td>SET</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Coastal Plain</td>
<td>CP</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Strzelecki Ranges</td>
<td>StR</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Western Otways</td>
<td>WOt</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Western Tasmania</td>
<td>WT</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Sub-race codes are used in Figure 1. SR Salmon River common garden site, TO Togari common garden site. doi:10.1371/journal.pone.0058416.0001

Genetic, environment and GxE effects on traits

To examine the variation in foliage chemistry across the two common environment trials, among sub-races and families within sub-races, as well as genetic-by-environment interactions, using only families from the native range (that is, not including the breeding programme seedlots) we fitted the following model with the PROC MIXED procedure of SAS:

\[
Y = \mu + \text{SITE} + \text{SUBRACE} + \text{SITE*SUBRACE} + \text{REP} + \text{FAM (SUBRACE)} + \text{SITE*FAM (SUBRACE)} + \text{RESIDUAL}
\]

whereby, \(Y\) is the phenotypic observation, \(\mu\) is the mean, SITE is the fixed site effect, SUBRACE is the fixed sub-race effect, SITE*SUBRACE is the fixed site by sub-race interaction effect, REP is the random replicate within site effect, FAM (SUBRACE) is the random family within sub-race effect, SITE*FAM (SUBRACE) is the random site by family within sub-race interaction effect, RESIDUAL is the residual. Fixed effects of SITE and SUBRACE, including the interaction term, were effectively our ‘treatment’ effects.

Discriminant analysis separating sub-races was conducted on a data set comprising the native-forest family means (calculated across sites) for the six chemical components using the PROC DISCRIM procedure of SAS. Mahalanobis distances among sub-races and their significance levels were calculated and summarised by UPGMA clustering undertaken with PROC CLUSTER.

Genetic parameters for each chemical constituent and DBH were estimated for each trial using the following univariate linear model:

\[
Y = \mu + \text{GROUP} + \text{SUBRACE} + \text{REP} + \text{FAM (SUBRACE)}_{\text{native}} + \text{FAM (SUBRACE)}_{\text{other}} + \text{RESIDUAL}_{\text{native}} + \text{RESIDUAL}_{\text{other}}
\]

where \(Y\) is the phenotypic observation, \(\mu\) is the mean, GROUP is the fixed population (either native forest progeny or other progeny from breeding programmes and commercial seed-orchards) effect, SUBRACE is the fixed native-forest sub-race effect, REP is the random replicate effect, FAM (SUBRACE)native is...
the random family effect (nested within sub-race) for the native forest progeny, \( FAM_{\text{after}} \) is the family effect for the non-native progeny, that is, seedlots from breeding programmes and commercial seed-orchards, \( RESIDUAL_{\text{after}} \) is the residuals for the native progeny, and \( RESIDUAL_{\text{other}} \) is the residuals for the non-native progeny. The utilisation of all available data (native forest progeny and other progeny from breeding programmes and commercial seed-orchards) in the analysis of individual sites allowed the effect of \( REP \) to be more precisely estimated, but only the results from native forest progeny are presented.

Additive genetic variance and heritability estimates

Due to mixed mating systems in eucalypts, additive variance \( (\sigma^2_a) \), was estimated from open pollinated family variances. The use of open pollinated seed to estimate additive variances is common in tree breeding programmes [26] where assumptions are made that there are no non-additive effects (maternal, dominance or epistasis) and that there is homogeneous outcrossing rates [37,38]. In \( E. \) globulus, maternal effects are minimal [39] and studies comparing genetic correlations between open pollinated and control pollinated estimates of additive genetic effects derived from a factorial crossing scheme for frost resistance [40], fungal pathogen damage [41] and other traits summarised in [42], show strong correlations between the open pollinated and controlled crosses. Consequently, we make the assumption of no non-additive effects for our analyses. For the native-forest families, the mean, additive variance \( (\sigma^2_a) \), phenotypic variance \( (\sigma^2_p) \), open-pollinated narrow-sense heritability \( (h^2_{\text{op}}) \) and coefficient of additive genetic variance \( (CV_a) \) [7] for each trait were estimated from univariate analysis as follows:

\[
\sigma^2_a = \frac{\sigma^2_f}{r}
\]

\[
\sigma^2_p = \sigma^2_f + \sigma^2_c
\]

\[
h^2_{\text{op}} = \frac{\sigma^2_a}{\sigma^2_p}
\]

\[
CV_a = 100 \times \frac{\sqrt{\sigma^2_a}}{\bar{x}}
\]

where \( \sigma^2_f \) is the family within sub-race variance [from \( FAM \) (\( SUBRACE \))]; \( \sigma^2_c \) is the residual variance [from \( RESIDUAL_{\text{after}} \)]; \( r \) is the coefficient of relationship, fixed to equal 0.4 to account for an assumed selling rate of 30% [43]; and \( \bar{x} \) is the trial mean. Standard errors of parameters were estimated from the average information matrix, using a standard truncated Taylor series approximation [34]. A ‘one-tailed’ likelihood ratio test was used to test whether the replicate and family within sub-race variance for each trait and site was significantly greater than zero [34]. The replicate term was dropped from the linear model for traits/sites for which the replicate variance was not significantly different from zero.

Phenotypic and genetic trait correlations

Bivariate analyses were undertaken for each trait to estimate pair-wise inter-site within sub-race family correlations (which were assumed to be equivalent to additive genetic correlations [44]). The model fitted included the same terms as the univariate model and treated each trait at different sites as a different variate with independent \( POPN \), \( SUBRACE \), \( REP \) and \( RESIDUAL \) effects. Furthermore, for each pairwise combination of traits, a multiple-site multiple-trait (i.e. four variates: two traits by two sites) analysis was conducted to estimate inter-trait genetic correlations. The multivariate models used for this purpose combined the aforementioned bivariate models and accounted for covariance between the \( REP \), \( FAM \) and \( RESIDUAL \) effects for traits within sites. Two-tailed likelihood ratio tests (LRTs) were conducted to test if inter-trait genetic correlations were heterogeneous across sites and to test whether the inter-trait and inter-site genetic correlations were significantly different from zero [45]. One-tailed LRTs were conducted to test genetic correlations against one and minus one. Sub-race correlations were estimated by fitting equivalent models with sub-race fitted as a random, rather than a fixed, term. Phenotypic correlations within sites were estimated in the same way as family correlations by removing all trial design (i.e. \( REP \)) and genetic terms (i.e. \( POPN \), \( SUBRACE \) and \( FAM \)) from the model.

Adaptive divergence amongst sub-races

To estimate the quantitative inbreeding coefficient \( (Q_{ST}) \) [46] for native populations, \( SUBRACE \) was fitted as random term in univariate analyses and \( Q_{ST} \) calculated as follows:

\[
Q_{ST} = \frac{\sigma^2_Q}{\sigma^2_a + 2\sigma^2_c}
\]

where \( \sigma^2_Q \) is the sub-race variance and \( \sigma^2_a \) is the additive variance. Following Dutkowski & Potts [47], ‘one-tailed’ likelihood ratio tests [34] were used, to test whether \( Q_{ST} \) estimates were significantly greater than a published estimate of the mean \( F_{ST} (0.09) \) which was based on eight microsatellite markers and used a similar suite of \( E. \) globulus sub-races [48]. More specifically, 10 of the 12 populations used in this current study were used in the \( FST \) study (where \( St \) Helens in this current study is named Humbug Hill in the \( FST \) study, Flinders Island in the current study is named north Furneaux in the \( FST \) study and finally, Southern Furneaux in the current study is named Cape Barren in the \( FST \) study). We similarly tested whether the \( Q_{ST} \) estimates were significantly greater than the maximum of the \( F_{ST} \) estimate for any microsatellite locus published for this species (0.201; see supplementary material 2 in [49]). \( F_{ST} \) is an estimate of the population-level inbreeding coefficient for neutral markers, which is assumed to reflect variation brought about by isolation and drift [46,48]. Assuming comparable mutation rates, significant deviation of \( Q_{ST} \) from \( F_{ST} \) is taken as indicative of stabilising \( (Q_{ST} < F_{ST}) \) or divergent \( (Q_{ST} > F_{ST}) \) selection having impacted on the genetic architecture of the trait [46,50]. In our case we were specifically interested in testing for evidence of divergent selection and thus used one-tailed tests.

Results

Significant differences among sub-races were evident for all chemical constituents (Table 2), with some sub-races containing twice the concentration of the chemical constituents as other sub-races (Figure 1). St Helens had the lowest levels of cineole, total oils and sideroxylonal while at the other extreme, the nearby sub-race, Flinders Island, had the highest level of these constituents (Figure 1). This pattern was reflected in the sub-race positive
### Table 2. Results of mixed model analysis of native-forest families examining site, genetic (sub-race and family within sub-race) and interactive effects for each chemical trait and diameter at breast height (DBH).

<table>
<thead>
<tr>
<th></th>
<th>sideroxylonal A</th>
<th>macrocarpal G</th>
<th>total oils</th>
<th>cineole</th>
<th>total phenolics</th>
<th>nitrogen</th>
<th>DBH</th>
<th>DBH#</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Site¹</td>
<td>38.3***</td>
<td>44.2***</td>
<td>122.1***</td>
<td>53.0***</td>
<td>26.8***</td>
<td>21.0***</td>
<td>1.0</td>
<td>4.8*</td>
</tr>
<tr>
<td>Sub-race²</td>
<td>11.1***</td>
<td>10.0***</td>
<td>4.3***</td>
<td>8.5***</td>
<td>6.1***</td>
<td>3.5**</td>
<td>4.7**</td>
<td>6.8***</td>
</tr>
<tr>
<td>Site*Sub-race²</td>
<td>0.9</td>
<td>0.7</td>
<td>0.5</td>
<td>0.7</td>
<td>0.5</td>
<td>0.6</td>
<td>2.1*</td>
<td>7.3***</td>
</tr>
<tr>
<td><strong>Random</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rep(Site)</td>
<td>0.8</td>
<td>0.9</td>
<td>2.1*</td>
<td>0.9</td>
<td>1.6</td>
<td>2.0*</td>
<td>0.0</td>
<td>3.6***</td>
</tr>
<tr>
<td>Family(Sub-race)</td>
<td>2.4**</td>
<td>3.5***</td>
<td>1.2</td>
<td>2.3*</td>
<td>2.3*</td>
<td>1.5</td>
<td>0.4</td>
<td>5.7***</td>
</tr>
<tr>
<td>Site*Family(Sub-race)</td>
<td>0.6</td>
<td>3.0</td>
<td>0.2</td>
<td>3.0</td>
<td>0.4</td>
<td>3.0</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>Residual</td>
<td>17.6***</td>
<td>18.0***</td>
<td>17.6***</td>
<td>18.0***</td>
<td>17.4***</td>
<td>18.1***</td>
<td>17.4***</td>
<td>45.1***</td>
</tr>
</tbody>
</table>

F (fixed terms) and Z (random terms) values and significance levels are shown.

|¹ numerator & denominator DF: 1, 33, except DBH#: 1,34 |
|² numerator & denominator DF: 11, 37, except DBH#: 12,113 |
|DBH analysis of the full dataset (not just trees with chemistry data). |
-*significant at P<0.05; **significant at P<0.01; *** significant at P<0.001.

doi:10.1371/journal.pone.0058416.t002

### Table 3. Phenotypic (rp), additive genetic (ra) and sub-race (rs) correlations between foliage chemistry, diameter at breast height (DBH) and browsing damage.

<table>
<thead>
<tr>
<th></th>
<th>sideroxylonal</th>
<th>macrocarpal</th>
<th>total oil</th>
<th>cineole</th>
<th>total phenolics</th>
<th>nitrogen</th>
<th>DBH</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>macrocarpal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rp</td>
<td>-0.23***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ra</td>
<td>-0.67**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs</td>
<td>-0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total oils</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rp</td>
<td>0.47***</td>
<td></td>
<td>0.45***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ra</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs</td>
<td>0.44</td>
<td></td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cineole</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rp</td>
<td>0.66***</td>
<td></td>
<td>0.32***</td>
<td>0.70***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ra</td>
<td>0.52</td>
<td></td>
<td>0.14</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs</td>
<td>0.87***</td>
<td></td>
<td>-0.11</td>
<td>0.77***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total phenolics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rp</td>
<td>-0.28***</td>
<td></td>
<td>-0.20***</td>
<td>-0.43***</td>
<td>-0.19***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ra</td>
<td>-0.40</td>
<td></td>
<td>0.50*</td>
<td>NA</td>
<td>0.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs</td>
<td>-0.04</td>
<td></td>
<td>-0.14</td>
<td>-0.49</td>
<td>-0.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nitrogen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rp</td>
<td>0.35***</td>
<td></td>
<td>0.04</td>
<td>0.22***</td>
<td>0.07*</td>
<td>-0.48***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ra</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs</td>
<td>-0.03</td>
<td></td>
<td>0.04</td>
<td>0.61</td>
<td>0.08</td>
<td>-0.92***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DBH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TO rp</td>
<td>0.11*</td>
<td>0.18***</td>
<td>0.24***</td>
<td>0.19***</td>
<td>-0.33***</td>
<td>0.21***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR rp</td>
<td>0.13*</td>
<td>0.20***</td>
<td>0.26***</td>
<td>0.16***</td>
<td>-0.33***</td>
<td>0.14***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TO ra</td>
<td>0.25</td>
<td>0.23</td>
<td>NA</td>
<td>0.26</td>
<td>-0.31</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR ra</td>
<td>0.17</td>
<td>0.02</td>
<td>NA</td>
<td>0.48</td>
<td>0.20</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TO rs</td>
<td>0.51</td>
<td>0.22</td>
<td>0.91*</td>
<td>0.77**</td>
<td>-0.33</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR rs</td>
<td>0.01</td>
<td>0.57</td>
<td>0.90*</td>
<td>0.16</td>
<td>-0.77*</td>
<td>0.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>browse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rp</td>
<td>-0.07</td>
<td>-0.26***</td>
<td>-0.26***</td>
<td>-0.27***</td>
<td>0.15*</td>
<td>0.14*</td>
<td>-0.30***</td>
<td></td>
</tr>
<tr>
<td>ra</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs</td>
<td>-0.42</td>
<td>-0.38</td>
<td>-0.57</td>
<td>-0.73*</td>
<td>-0.43</td>
<td>0.79</td>
<td>0.70*</td>
<td></td>
</tr>
</tbody>
</table>

For Togari field site, SR Salmon River field site.

Correlations for DBH were presented separately for Togari and Salmon River field sites due to genotype by environment interaction at the sub-race level.

NA = no additive genetic variance

*significant at P<0.05; **significant at P<0.01; *** significant at P<0.001.

. = significance could not be determined using a likelihood ratio test.

Correlations with browse damage are using data from Togari common garden trail only.

doi:10.1371/journal.pone.0058416.t003
correlations between cineole and sideroxylonal (Table 3). Several patterns in chemical concentration emerge from Figure 1. Macrocarpal G concentration (Figure 1b) appears highest in the western Victorian sub-races (the Otways), lowest in the eastern Victorian sub-races and intermediate in the Tasmanian populations. Interestingly, patterns for sideroxylonal (Figure 1a) show that the Flinders Island sub-race is similar in concentration to the eastern Victorian sub-races, Southern Furneaux is intermediate and St. Helens on the north-east of Tasmania is lowest, showing a steep cline to north-eastern Tasmania. Additionally, mainland sub-races are lower in concentration of total phenolics compared to Tasmanian sub-races and the opposite pattern is evident for nitrogen. These patterns were reflected in the sub-race inter-trait correlations with a negative correlation between total phenolics and nitrogen (Table 3).

Cluster analysis based on the Mahalanobis distances (Table S1) amongst sub-races resulted in five major sub-race groupings (Figure 2): St Helens and Flinders Island separated independently (although Flinders Island grouped closely to the south-eastern Victoria sub-races), Eastern and Western Otways grouped together, Coastal Plain and Strezleki Ranges grouped together, and the remaining Tasmanian sub-races formed the fifth group. Discriminant analysis identified three significant canonical variates (CVs); CV1 accounted for 51.7% of the variation, CV2 25.8% and CV3 16.4% of the variation, respectively. When CVs were mapped as per Figure 1 (CV maps not shown), CV1 mirrored the distribution of sideroxylonal, CV2 mirrored the distribution of total oils and CV3 mirrored the distribution of nitrogen and total phenolics, indicating that these CVs were strongly influenced by these compounds.

As well as significant among sub-race variation in chemical constituents, there was also significant family within sub-race variation (Table 2) and significant additive genetic variance estimates (Table 4) for the majority of compounds, except for total oils and nitrogen. Open-pollinated narrow sense heritability estimates ranged from 0.07 for nitrogen to 0.48 for macrocarpal. Heritability estimates were not significantly different between sites ($p>0.05$). The ten Qst estimates for the foliar PSMs ranged from 0.20 to 0.71 and averaged 0.36±0.05. This average was
Hierarchical genetic variation in PSMs and evidence of selection

A major aim of this paper was to estimate genetic parameters (heritabilities and coefficients of additive genetic variance) that inform us of how responsive PSM traits may be to selection [7,26]. Previous heritability estimates of the ecologically important FPCs in *E. globulus* have been based on broad sense heritability estimates (macrocarpal G 0.51 ± 0.07, sideroxylonal A 0.79 ± 0.04 [51]) but here we report narrow sense heritabilities; which reflect the expression of many PSMs resides both among and within populations of *E. globulus*, (ii) the expression of this genetic variation is stable across environments, and (iii) that diversifying selection is likely to have acted on some PSMs.

**Discussion**

In this paper we link chemical variation with evolutionary dynamics and ecological interactions across different levels of organisation to demonstrate that (i) significant genetic variation in the expression of many PSMs resides both among and within populations of *E. globulus*, (ii) the expression of this genetic variation is stable across environments, and (iii) that diversifying selection is likely to have acted on some PSMs.
slightly lower for sideroxylonal A (0.33±0.14 to 0.24±0.12, respectively). Narrow sense estimates for sideroxylonal have also been reported in another eucalypt species where estimates were high (e.g. 0.39±0.12−0.62±0.12[5]). Heritability estimates obviously vary depending on species, populations and environments [17] but what these estimates tell us from these few studies is that these important PSMs are under relatively strong genetic control and there is adequate additive genetic variance within populations for selection. The later is particularly the case for the two FPC compounds, where the coefficient of additive genetic variance (CAV; >15%) is atypically high compared with other traits in E. globulus (present study and [52]) other eucalypts [53] and other taxa [7]. This clearly shows these FPC compounds have substantial genetic variation available for exploitation, either through natural or artificial selection, in this species.

Indeed, our examination of differences in the Qst and Fst estimates suggest that sub-race divergence in concentration of foliar PSMs have already evolved through selection. Differences in quantitative trait and neutral molecular estimates have been calculated in a number of plant species to indicate patterns of selection for a range of life history traits [54,55]. Here, our average Qst estimates for foliar PSMs is significantly higher than the average of Fst estimates for putatively neutral microsatellite loci. For two PSMs, sideroxylonal and cineole, the Qst is significantly higher than the published maximum Fst estimate for putatively neutral microsatellite loci [48], and are amongst the highest Qst values yet estimated for E. globulus. While other factors, such as strong genetic correlations amongst traits and correlated responses to selection cannot be dismissed [56], these results argue that natural selection has contributed to sub-race divergence in PSMs in E. globulus, and identifies the FPCs (particularly sideroxylonal) and cineole as candidates for traits whose genetic architecture has been shaped by divergent selection [1,48].

**Possible agents of selection on PSMs in E. globulus**

Plant secondary metabolites in E. globulus may play a role in numerous ecological functions including defence against mammalian herbivores and fire promotion [57,58]. We have previously suggested that some PSMs in E. globulus are under selection by mammalian herbivores because mammalian browsing preferences differ amongst E. globulus populations and these differences can be explained, in part, by variation in PSMs among those populations, particularly the FPCs, sideroxylonal and macrocarpal [18,19]. However, this is the first time that Qst estimates have been calculated for PSMs in this dominant tree species providing an independent line of evidence that these traits have been under divergent selection. Here, we also link these estimates with observed negative phenotypic correlations between field browsing by mammalian herbivores, macrocarpal G and cineole (however, negative correlations between browsing and sideroxylonal were not significant in this study). There were also negative phenotypic and additive genetic correlations between field browsing and a surrogate of fitness, stem diameter growth (DBH). In this present study, trees that received less browsing had higher levels of macrocarpal G and cineole and had larger diameters where browsing reduces growth rates as opposed to browsers selecting smaller trees [39]. In support of this surrogate fitness measure, longer-term genetic studies in similar competitive plantings have shown E. globulus survival is size dependent [52]. Given the relationships with browsing, FPCs (e.g. [18,19]; and this present study for macrocarpal) and growth, the idea that cineole acts as an olfactory cue to the FPCs [60] and the genetic data presented in this present study, there is clear evidence that PSMs in E. globulus are under selection and that that mammalian herbivores may play a role. The correlated nature of the compounds [19], however, makes it less clear exactly which traits may be under direct selection.

Obviously, fitness estimates are difficult to obtain in a long lived tree species and hence surrogates of fitness (such as DBH and reproductive output at a snapshot in time) are often our best option. While we have only assessed DBH in this current study, we have shown that repeated browsing by mammalian herbivores greatly reduces individual plant reproductive fitness (number of buds and capsules) at 10 years of age compared to those plants that were not browsed [61]. In this current paper, there is also little evidence of a trade-off between PSM production and growth [62,63], although we cannot discount costs at different life stages of the plant [64]. Most phenotypic correlations between PSMs and DBH were significant and positive (apart from total phenolics) and we did not detect any significant negative correlations at the sub-race or additive levels.

**Population divergence and geographic patterns of PSMs in E. globulus**

Broad-scale population divergence in the expression of PSMs in this study appear to follow some clear geographic patterns. The two mainland Australian regions, the south-western Victorian sub-races (Eastern and Western Otway) and the south-eastern Victorian sub-races (Skerzelecki and Coastal Plains) show divergent patterns of PSM expression. If one region is high in concentration of PSM then the other region is low (evident for sideroxylonal A, macrocarpal G and total oils). These two distinct groupings are also reflected in the UPGMA clustering of sub-races based on variation in PSMs presented in this current paper and also in Wallis et al [22] where the two Victorian regions were shown to be dominated by different ‘chemotypes’. Molecular studies have shown that these two mainland regions also differentiated on nuclear microsatellite markers [48,49], chloroplast DNA haplotype frequencies (Freeman et al, 2001) and recent diversity arrays technology markers (DArT) [63]. Hence, genetic differentiation in these PSM traits also appears to be associated with reduced gene flow between these regions.

Another clear pattern is the steep latitudinal cline in most PSMs (sideroxylonal A, total oils, cineole and total phenolics) from Flinders Island, Southern Furneaux through to the north-eastern Tasmanian sub-race of St Helens and where Flinders island groups more closely to the south-eastern Victorian sub-races. Clinal patterns in quantitative traits have become a point of interest in recent literature due to probable effects of large scale environmental variation (e.g. climate change effects) on local adaptation of populations [66]. Our clinal trends in chemistry are consistent with that reported across the island group by Wallis et al [22] which appears to coincide with a barrier to seed dispersal suggested by the north-eastern Tasmanian populations lacking a specific chloroplast haplotype which is common in eastern Victoria and present on Flinders Island and the Southern Furneaux [67]. However, while there may be a barrier to seed dispersal, there is evidence of pollen flow [68] supporting the steep cline evident in this study.

Finally, based on a chloroplast DNA study, Freeman et al. [67] also postulated a recent ‘western-link’ between the mainland and Tasmania, from the Otways, through King Island and down the west coast and moving up the east coast of Tasmania. The genetic-based clinal differentiation in PSMs reported in Wallis et al. [22] supported this pattern and suggests a clinal loss of plant PSM defences along this gradient. This appears to have culminated in the evolution of the least defended north-eastern Tasmania populations [18,19]. While our study argues for the role of
disruptive selection and marsupial herbivory in shaping these subrace patterns of differentiation [but to date we have no evidence of differential selection pressure], the steep cline across the Furneaux Islands potentially represents a dynamic zone of secondary contact and pollen-mediated gene flow between the highly defended, eastern mainland and least defended, north-eastern Tasmanian gene pools [22].

Lack of G × E for PSMs in E. globulus

Variability in the genetic expression of traits across environments provides an indication of the relative strength of genetics in affecting ecologically important traits [69,69,70]. This concept is important particularly for plant species that show extended genetic effects and community consequences [71]. To date, only a few studies have examined the plasticity of genetic based variation of PSMs in eucalypts across environments using common pedigrees [5,72,73] and few studies have examined G × E interactions of plant chemical defences using a genetic hierarchy in any plant/herbivore system [5]. Here we provide evidence that both among and within population genetic variation in PSMs in E. globulus is stable across sites. We did not detect a genotype by environment interaction (G × E) for any of the chemical traits, despite highly significant genetic and site main effects. Previously in nutrient trials, we detected a G × E on mammalian herbivore responses to E. globulus in a captive feeding trial, but did not detect a G × E with any of the PSMs we assessed in the seedlings. Additionally, Andrew et al. [5] demonstrated a G × E interaction on sideroxylonal in Eucalyptus tricarpa at the population level, but not within populations. In some recent work, we manipulated CO₂ levels in the glasshouse with different populations of two eucalypt species, E. globulus and E. pauciflora, and found that CO₂ had minimal effect on the expression of PSMs and no G × E was detected [73].

What is becoming apparent in eucalypt systems is that while the levels of PSMs may change with environment, the relative differences amongst genotypes remains stable. Low or absent G × E for eucalypt PSMs has important ecological and applied management implications. If these genetic based traits that influence community phenotypes are stable across environments, the role of plant genotype in structuring communities as observed in E. globulus [20], is strengthened and these genotypic differences may be relatively stable under global environmental change such as increasing atmospheric CO₂. Additionally, as many eucalypt species are important commercial forestry species (E. globulus is the most widely planted hardwood species in temperate regions worldwide) the knowledge that genotypes may be stable in the expression of PSMs across different plantation sites give confidence to tree growers for artificial selection for these traits in order to reduce losses from pest damage [74].

Supporting Information

Table S1 Table of pair-wise Mahalanobis distances amongst E. globulus sub-races and their significance.

The Mahalanobis distances were calculated using the PROC DISCRIM procedure of SAS and data set comprising the native-forest family means (calculated across sites) for the six chemical components analysed.

(DOC)

Acknowledgments

We would like to thank Jonathan Humphreys for assistance with the NIRS analysis, and Hugh Fitzgerald and Paul Tilyard for assistance with field work. We thank anonymous reviewers for comments on the manuscript.

We would like to thank Forestry Tasmania for access to field sites and the CRC for Forestry for support.

Author Contributions

Conceived and designed the experiments: JOR-W BP DW. Performed the experiments: JOR-W AM NG-D. Analyzed the data: AM BP MH. Wrote the paper: JOR-W AM BP MH.

References